

Document downloaded from:

<http://hdl.handle.net/10251/120245>

This paper must be cited as:

Cabrera Rochera, E.; Estruch-Juan, ME.; Molinos-Senante, M. (2018). Adequacy of DEA as a regulatory tool in the water sector. The impact of data uncertainty. *Environmental Science & Policy*. 85:155-162. <https://doi.org/10.1016/j.envsci.2018.03.028>



The final publication is available at

<https://doi.org/10.1016/j.envsci.2018.03.028>

Copyright Elsevier

Additional Information

Adequacy of DEA as a regulatory tool in the water sector. The impact of data uncertainty.

E. Cabrera Rochera¹, E. Estruch-Juan², M. Molinos-Senante^{3,4,5}

¹ ITA, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain. Corresponding author (E-mail: qcabrera@ita.upv.es)

²ITA, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain (E-mail: maesjua1@ita.upv.es)

³ Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, Chile (E-mail: mmolinos@uc.cl)

^{4b} Instituto de Estudios Urbanos y Territoriales, Pontificia Universidad Católica de Chile, El Comendador 1916, Santiago, Chile.

⁵ Centro de Desarrollo Urbano Sustentable CONICYT/FONDAP/15110020, Av. Vicuña Mackenna 4860, Santiago, Chile.

Abstract

The regulation of water services shares many similarities with that of other utilities such as electricity or telecommunications. As a result, similar methods are often used by regulators to assess the efficiency of companies in those sectors. Data Envelopment Analysis (DEA) is one of those widely applied methods. This paper aims to determine the adequacy of DEA as a regulatory tool for urban water services, with a special focus on the quality of the available data. In order to obtain useful conclusions, two DEA simulations were performed with audited data from 194 water utilities, officially made available by the Portuguese water regulatory authority (ERSAR). Both simulations will demonstrate that the inherent inaccuracies found in some of the key data provided by water utilities represent a significant obstacle to obtain meaningful results with the DEA technique. This could represent a paradigm shift for some of the regulatory authorities currently using DEA or similar techniques, as the complexity of the method does not seem to be justified by a better analysis of the comparative performance of the different services.

Keywords:

Data envelopment analysis (DEA), Uncertainty, regulation, yardstick competition, benchmarking, water services

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1. Introduction

Urban water services are usually provided in a natural monopoly market, for the entire demand in the supply area can be satisfied at a lowest cost by one firm

rather than by two or more (Posner, 1969). In the case of water services (like in most other public services described as natural monopolies), the key factor for this monopolistic environment is that a large fraction of the total costs consists of sunk capital costs (Joskow, 2007). In the case of water services, the capital costs correspond to the networks usually needed to deliver those services.

Some authors consider that the natural monopoly is an economic fiction, and that the public utilities were in vigorous competition until government-sanctioned monopolies were granted (DiLorenzo, 1996). However, even considering this arguable contention to the natural monopoly concept, water has probably entered in a category of its own since the United Nations declared in 2010 safe and clean drinking water and sanitation a human right (United Nations, 2010).

Regulation is the response of governments and public administration to protect the interests of the users in these natural monopolies. The Lisbon Charter (IWA, 2015) lists among the responsibilities of regulatory authorities for water services to supervise tariffs, oversee and promote a suitable quality of service, ensure the protection of consumer rights, contribute to fair and open competition between service providers, collect analyse and disseminate accurate information and provide incentives for improvement of the services.

Regulators have been addressing these responsibilities, at least in part, by collecting key data and calculating performance indicators to assess the performance of the different regulated services. Some of these regulators have collected, processed (and sometimes made public) this information to understand the performance of each operator, and used it in a qualitative manner to address their regulatory duties.

However, other regulators (starting in the early 90s with OFWAT, the economic regulatory authority for England and Wales) decided to compare the performance of the different services in what was coined as “yardstick competition”. In the original reference, Shleifer (1985) defined it as *“the simultaneous regulation of identical or similar firms. Under this scheme, the rewards of a given firm depend on its standing vis-à-vis a shadow firm, constructed from suitably averaging the choices of other firms in the group”*. In the same paper, Shleifer presented the use of regression techniques to address the heterogeneity of the firms (in terms of context, and therefore of cost drivers). Shortly afterwards, yardstick competition became an integral part of the regulation scheme of OFWAT. However, even at this early stage, it was clear that the practical implementation of Shleifer theory was far from easy. Sawkins (1995) acknowledged these struggles: *“In total then, even this very limited form of yardstick competition, imperfect and at an early stage of development, was delivering results”*.

Following the leading steps of OFWAT, the world has witnessed in the past 20 years many attempts of overcoming one of the difficulties originally outlined by Shleifer: the exogenous characteristics of utilities that drive costs. This issue has been addressed with models of increasing complexity, trying to determine the cost drivers and therefore establish a cost function that will enable to assess the degree of efficiency of a service.

The literature presents a wealth of efficiency studies in the water sector using frontier models. Byrnes et al. (2010) for instance, use regulatory data in New South Wales and Victoria to determine the relative economic efficiency of 52 water utilities. In their literature review, they cite 20 additional studies. Non-parametric methods have also been widely employed to evaluate and compare the performance of water companies (Worthington, 2014). Within this approach, the most used method is data envelopment analysis (DEA) (Romano and Guerrini, 2011). Unlike parametric methods, DEA does not require to define the functional form of the production frontier (De Witte and Marques, 2010).

DEA has been widely applied to evaluate the performance of water companies. However, DEA is a deterministic method and therefore cannot deal with imprecise data or provide information about uncertainty (Kao and Liu, 2014). This means that traditional DEA models (Charnes et al., 1978; Banker et al., 1984) do not permit stochastic variations and uncertainty in data, and require that the exact values of all inputs and outputs are known. However, these assumptions may be faulty, since some data cannot be measured accurately enough in practice (Eslami et al., 2012). As a matter of fact, uncertainty naturally exists in data that are either collected, monitored or recorded in water services.

To overcome this problem and to take uncertainty into account, several extensions to the traditional DEA models have been proposed, such as Monte Carlo simulation, the α -level based approach, chance constraint, bootstrapping, fuzzy ranking, and DEA-tolerance. In the case of water utilities, these methodologies have barely been used with the exception of De Witte and Marques (2010a) who applied *orden-m* method to incorporate environmental variables into water companies' efficiency assessment and De Witte and Marques (2010b), Ananda (2014), See (2015) and Molinos-Senante et al. (2018) who employed a double-bootstrap DEA approach to compute bias-corrected efficiency scores. Each of these methodological approaches has advantages and shortcomings. However, Bonilla et al. (2004) showed that the DEA-tolerance method is simpler and faster than the bootstrapping approach, and leads to similar results. Moreover, Dong et al. (2017) noted that the DEA-tolerance approach is less subjective than the fuzzy approach since it does not need the fuzzy sets of variables to be defined for units. Furthermore, the DEA-

tolerance approach can be combined with the system of indicators proposed by Boscá et al. (2011) that allows units to be benchmarked in an uncertain context.

Given the nature of water services, the information in the water sector is quite poor. Most assets are buried underground and they may have been there for decades or even over a century. Water services are also more complex from a quality of service perspective (Cabrera & Cabrera Jr. 2016) and may require a greater number of variables to be taken into account. Some of those variables are not easy to measure accurately at a reasonable cost (e.g. flow rates, pressures...). Additionally, a significant portion of the delivered product (water) is lost through leaks, and the importance of this loss can only be estimated (and not measured). Remote sensing technologies are more complex to deploy for water services in comparison with the electricity or telecommunications services. The water industry is well aware of this limitation and both the ISO standards International Water Association strongly recommends to record the quality of input data in any performance assessment system (ISO 24510:2007, ISO 24511:2007, ISO 24512:2007, Alegre et al., 2016).

In spite of data inaccuracy, most examples in the literature using DEA to evaluate the efficiency of water companies focus on the frontier model to be used, the input-output model or the variables to adopt as inputs or outputs without taking care of data uncertainty. Thus, data quality is seldom covered by DEA studies and the sensitivity of the methodology to poor data has not been studied in sufficient detail. This is surprising as data quality in the water industry is notoriously deficient (Alegre et al., 2016).

This paper aims to determine whether the quality of data available in the water industry (and its uncertainty) suffices to obtain meaningful conclusions when trying to compare the efficiency of different operators using DEA. In order to do so, the technique has been used with real data from 194 utilities from a regulated environment (Portugal, where ERSAR is the regulator). This dataset has a significant characteristic: for all data, and following ERSAR's requirements, all utilities recorded the uncertainty associated with each individual number. The data figures and their uncertainty can be considered reliable, as ERSAR audits all data submissions from utilities.

2. Methodology

The role of uncertainty of data in DEA models is essential, because the conclusions derived from performance analyses are highly sensitive to errors in data (Bhardwaj et al., 2018). To integrate uncertainty in performance assessment of water companies, the DEA-tolerance methodology was applied. Unlike traditional DEA models, the DEA-tolerance approach captures

uncertainty by constructing intervals for data (Dyson and Shale, 2010). Thus, it provides information on how sensitive water companies are to performance, with respect to changes in inputs and outputs by considering several scenarios for each water company.

The definition of the tolerance values for inputs and outputs is an essential step in applying the DEA-tolerance model. In this case study, tolerances for each water company and variable were based on the degree of uncertainty declared by water companies themselves for every single variable (see Section 3 for additional information). Thus, asymmetric and variable tolerances were defined for each piece of data.

Given $k = 1, 2, \dots, n$, water companies with each one using a vector of M inputs $x_k = (x_{1k}, x_{2k}, \dots, x_{Mk})$ to produce a vector of S outputs $y_k = (y_{1k}, y_{2k}, \dots, y_{Sk})$, the tolerances defined are non-negative scalar values and express the positive and negative changes in the values of the inputs and outputs as follows:

$$\text{Tolerance for inputs: } \alpha_{ik} = x_{ik}r_{ik} \text{ and } \alpha'_{ik} = x_{ik}z_{ik} \quad (1)$$

$$\text{Tolerance for outputs: } \beta_{rk} = y_{rk}s_{rk} \text{ and } \beta'_{rk} = y_{rk}t_{rk}$$

where:

r_{ik} , z_{ik} , s_{rk} , and t_{rk} are percentages of negative and positive deviation from the original values for inputs and outputs, and may range from [0 – 100].

According to the tolerance values defined, the values of the inputs and outputs should be within the following range:

$$\begin{aligned} x_{ik} &\in [x_{ik} - \alpha_{ik}, x_{ik} + \alpha'_{ik}] \\ y_{rk} &\in [y_{rk} - \beta_{rk}, y_{rk} + \beta'_{rk}] \end{aligned} \quad (2)$$

Equation (2) shows that there is a breadth of possible combinations of inputs and outputs. Thus, scenarios in which the performance indexes are computed should be selected. Following Molinos-Senante et al. (2016), 81 scenarios were simulated for each water company which allowed the maximum (best case) and minimum (worst case) performance index to be obtained, in addition to the original and mean values. These 81 scenarios result from calculating 3^4 alternatives, corresponding to three situations for each utility: i) favorable; ii) unfavorable, and iii) original, with four possible input and output sets for each one of them: i) inputs for the analysed Decision Making Unit (DMU); ii) outputs for the analysed DMU; iii) inputs for the remaining DMUs, and iv) outputs for the remaining DMUs. The best-case scenario (the best for the water company evaluated) involves considering the lowest values for inputs and the highest

values for outputs. By contrast, the worst-case scenario involves the opposite situation.

Once tolerance values were defined, performance indexes were calculated for each water company for all 81 scenarios. Following previous studies (Marques et al., 2014; Guerrini et al., 2015; Molinos-Senante et al., 2016), a DEA approach based on assumptions of variable returns to scale (VRS) and minimization orientation (Wang et al., 2012) (Eq. (3)) was employed to compute a performance index for each water company evaluated. The measure of performance θ is obtained, according to the DEA-VRS model, by solving the following linear programming problem for each water company k_0 , for each of the 81 defined scenarios:

$$\begin{aligned}
 & \text{Min } \theta \\
 & \text{s. t.} \\
 & \sum_{k=1}^n \lambda_k x_{ik} \leq \theta x_{ik_0} & 1 \leq i \leq M \\
 & \sum_{k=1}^n \lambda_k y_{rk} \geq y_{rk_0} & 1 \leq r \leq S \\
 & \sum_{k=1}^n \lambda_k = 1 \\
 & \lambda_k \geq 0 & 1 \leq k \leq n
 \end{aligned} \tag{3}$$

where λ_k is a vector of intensity. The measure of performance θ ranges between 0 and 1. A water company is considered efficient if $\theta = 1$, since it means that the water company is located on the efficient frontier. In contrast, an efficiency lower than 1 indicates that the performance of the company can improve and that the utility is therefore inefficient.

Under the best and worst-case scenarios, the maximum and minimum performance indexes were obtained for each water company evaluated, allowing analysing the impact of data uncertainty on the performance of water companies.

3. Sample description

The described methodology was applied to a dataset collected by the Portuguese Water and Waste Services Regulation Authority (ERSAR) in 2015 (ERSAR, 2015). The selection of the data was motivated by several reasons. In first place, all data variables were accompanied by an uncertainty value and all data and uncertainty values provided by the utilities were audited by ERSAR. Secondly, data are publicly available through the ERSAR website. Finally, ERSAR was considered to be a reliable source of data, with over 15 years of experience in data collection and publication of results (ERSAR, 2017), and a reputation as a regulator of excellence.

The DEA model used was designed to estimate the efficiency of water supply services. For this reason, utilities providing water supply services were selected from the ERSAR database (including multi-utilities), excluding those that solely provide wastewater or solid waste services. For this reason, special care was taken in the selection of variables in order to avoid the use variables that might be influenced by performance of the other services.

It is of great importance when assessing the efficiency of water supply utilities to consider the quality of service provided. Maintaining a high quality of the service has a clear impact on costs (Picazo-Tadeo et al., 2008) and, in consequence, efficiency results may be significantly biased if quality of service is not considered. Quality of service has been traditionally omitted from the efficiency models, as a deep analysis of most of the frontier efficiency models published applied to the water sector proves (Ferro et al., 2014; Walter et al., 2009). Even though the sector is slowly becoming aware of the impact that quality of service has on results (Picazo-Tadeo et al., 2008; Molinos-Senante et al., 2016; Pinto et al., 2016) a standard procedure has not been defined yet to consider it in frontier efficiency methods. In this research, quality of service influences both inputs and outputs.

All variables used as inputs and outputs in the DEA model are defined in detail in the ERSAR technical guides (ERSAR 2011). Following previous studies (See, 2015), the model considers four inputs: network length (km), total expenses (€/year), volume of real water losses (m³/year) and service interruptions (number/year). The last two inputs provide information on the quality of service. Water companies should minimize both variables to provide high-quality service and therefore, they were integrated in the model as inputs.

Total expenses were used instead of using operating and labour costs, which are variables typically considered in the water sector (Ferro et al. 2014). Total expenses include capital and operating expenses. This variable have not been extensively used for frontier efficiency methods in the water sector due to the disparity in the valuation of the capital expenses. However, in this study they can be included as inputs due to the fact that the sample belongs to a regulated environment and all utilities use the same definition for their calculation.

The selected outputs were: volume of produced water (m³/year) and number of households covered by the water service (No.). Both variables have been widely used in frontier efficiency studies as outputs (Ferro et al., 2014; Abbott & Cohen, 2009).

Concerning the number of utilities, from the 265 utilities providing water services in the ERSAR database, only those reporting uncertainty information for all variables used as inputs/outputs in the DEA model were considered. As a

consequence, only 194 were considered as DMUs in the efficiency model, discarding the remaining 70 due to incomplete data quality information.

The number of DMUs can be a limitation in a DEA model. A minimum number of DMUs is needed and this number is determined by the amount of inputs and outputs considered (Tupper & Resende 2004). This number can be calculated with “Cooper’s rule”: the number of DMUs to be evaluated in the model must be equal or larger than $\max\{m \times s; 3(m + s)\}$ where m is the number of inputs and s the number of outputs (Cooper et al. 2011). In this study, the 194 DMUs considered exceed by far the requested minimum number.

In the water sector, with the entire infrastructure buried and a significant amount of data estimated, reliable and accurate data is hard to find. For instance, considering water volumes, the error of a water meter can range, depending on the technology, between 0.2% and 5% (Arregui et al., 2007). However, these values are only valid for new, calibrated, properly installed and appropriately selected water meters (working in their designed flow rate range) (Arregui et al., 2007). These values increase with water meters not complying with these specifications and far larger errors can be found when volumes are estimated (unmetered consumption, real losses and commercial losses).

Data quality in the considered dataset was taken into account by means of uncertainty bands. These uncertainty bands reflect the confidence grade of the values submitted by the utility and were audited by ERSAR. Uncertainty was considered in 6 discrete bands (ERSAR, 2011). Table 1 displays the uncertainty bands considered.

The uncertainty associated to each input or output is different for each variable and utility. For instance, for service interruptions output, utility A can have an uncertainty band of 0-5% as they have an updated digital record system. Utility B uncertainty, ranges between 20-50% if they have a non-updated recording system in paper format.

3.1. Uncertainty band analysis

The number of utilities with data in each uncertainty band was analysed for both inputs and outputs. Results are summarized in Table 1. The most accurate variables, are volume of water produced and number of households with water service.

In the case of the volume of water produced, this volume is usually metered after the water treatment plant (WTP) with a bulk meter. Uncertainty in this case is driven by the accuracy of the metering system used, the age of the meter and whether it has been appropriately selected and calibrated. Optimum values

belong to the 0-5% band (Arregui et al., 2007; ISO 4064-1, 2014) and values out of this range would correspond to older, non-calibrated or not appropriately selected meters and for non-metered volumes (estimated by the utility). The Number of households with water service variable also presents lower uncertainties as this value is generally accurately known through to the billing system.

Knowledge of the network length is far less accurate. The main sources of inaccuracies in this value derive from the networks' age, the lack of proper records for network extensions in new urban developments and diversity of sources. Even though the average uncertainty is higher than in previous variables, it is maintained within reasonable levels.

Moving to total expenses, this is the second variable with fewer utilities in the 0-5% band. This higher uncertainty derives from the fact that some utilities did not have appropriate costing systems at the time the data were collected.

Finally, volume of water losses and number of interruptions present high levels of uncertainty. The volume of water losses is obtained from the water balance, with significant uncertainties potentially associated to measured volumes, with some estimated volumes also being part of the calculation.

Regarding the number of interruptions, 92% of values are located in the 0-20% uncertainty band. This value entirely depends on how utilities are able to record interruptions and their significant related data. In a regulated environment, this value should be quite accurate.

Table 1. Number of utilities belonging to each uncertainty band classified for each input and output

Uncertainty bands	Inputs				Outputs	
	Network length	Total expenses	Nº of interruptions	Volume of real water losses	Nº of households	Volume of water produced
0 - 5%	133	85	147	83	175	166
5 - 20%	45	60	31	69	15	17
20 - 50%	11	34	5	28	2	6
50 - 100%	5	8	9	11	1	5
100 - 300%	0	7	0	2	1	0
> 300%	0	0	2	1	0	0

4. Results and discussion

81 DEA simulation were performed for the entire sample of 194 utilities. Figure 1 represents the range of results for each utility. The DEA simulation with the

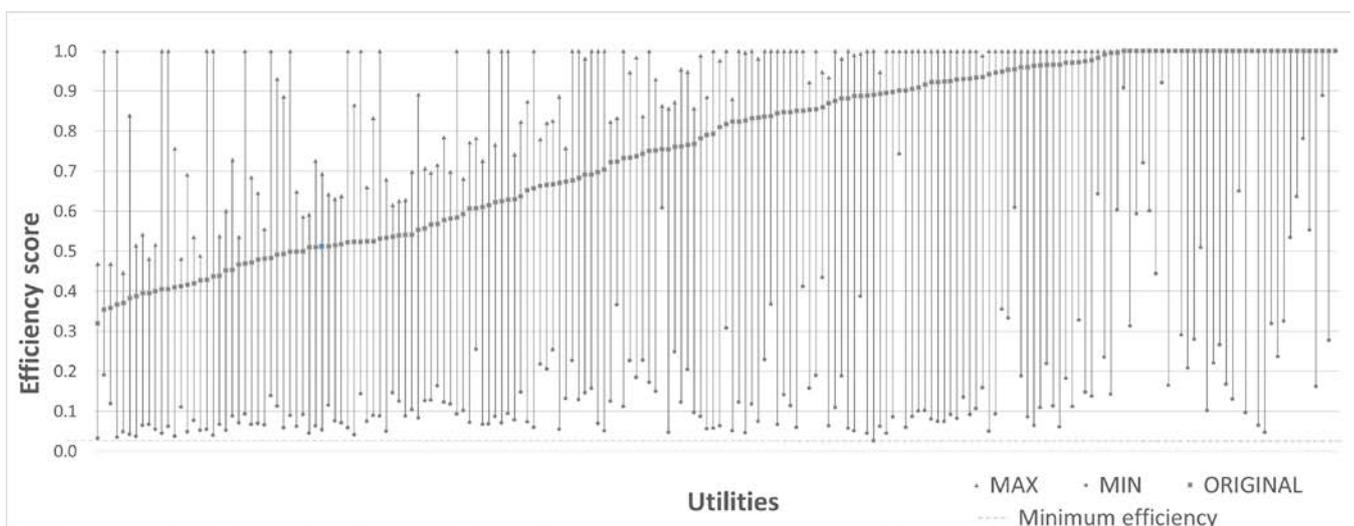
original data submitted to the regulator are represented in this figure by the series labelled as “Original” which is characterized by a square. Utilities are ordered by the original efficiency, as shown by the positive slope of this series. The maximum efficiency for each utility in the 81 simulations (its best result taking into account uncertainty in data) is represented by a triangle and is named “Max”. The minimum efficiency (worst result out of all scenarios) is represented by the circle and labelled as “Min”.

The vertical lines represent the efficiency amplitude (maximum-minimum efficiency score) obtained by each utility. In other words, the gap between the best and worst efficiencies estimated by the DEA simulation. The discrepancies in the results are very significant, and the average change in efficiency for the sample is 71%, with a maximum of 97% for one of the utilities. This result evidences the importance of data uncertainty in efficiency assessment of water companies.

This variability in results is derived from the uncertainty of input and output data declared by the utilities themselves and audited by the regulator. This amplitude is difficult to rationalize, as a 71% amplitude can turn an efficient utility, sitting at the frontier (those with a triangle at the top line, with an efficiency score of 1) into an inefficient one (0.29 efficiency score, or a 29% efficiency).

It is very important to note that these 81 scenarios do not correspond to possible alternatives in external or internal events or conditions, and simply address the inaccuracies of records or measuring devices. In other words, they all correspond to the current reality. An audited reality.

Figure 1: Maximum, minimum and original DEA values



Efficient utilities are those with an efficiency score of 1. In the simulation with the original submitted data, 34 utilities (17.5%) are fully efficient. These efficient utilities conform the frontier and represent the benchmark for the remaining utilities.

In the scenario with fewer utilities at the frontier, only 3 of them reach a score of 1, while in the scenario with a higher number of efficient utilities, 110 utilities conform the frontier. Considering these results, a regulator would have a very hard time separating the efficient competitors from those who are not using only these data and method.

DEA is often used to rank utilities against each other creating yardstick competition. The relative position of the utility in the ranking is used to trigger regulatory measures such as penalties, incentives, tariff increases, etc. In order to determine how data quality affected the relative position of utilities in a yardstick competition situation, the position in the ranking for each utility in all 81 simulations was obtained. The most efficient utilities are ranked at the top, while the last positions represent low efficiency.

Table 2 shows the variability of the ranking using four individual utilities (numbers 24, 106, 112 and 163) and five different DEA simulations (out of the 81), including the Original DEA simulation. The entire ranking is available as additional data for this paper.

Once again, the use of the results for regulatory purposes becomes impossible. Depending on the considered simulation (all within the uncertainties submitted by the utilities), the top utility in the original simulation (number 163) can go as low in the ranking as position 184. All four utilities are subject to changes in their ranking (depending on the combinations of possible data) of up to 192 positions (utility 106 experiments the smaller variation of the four, with 95 positions).

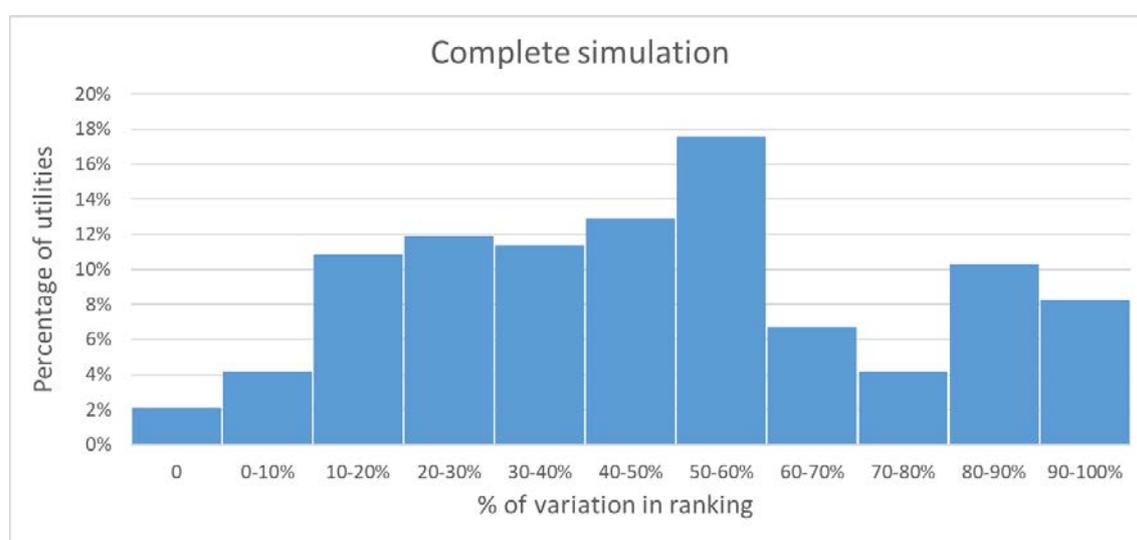
Table 2. Example of utilities variation in the ranking

ID Utility	ORIGINAL DEA Position	Simulation 6		Simulation 8		Simulation 18		Simulation 34	
	Ranking position	Ranking position	Δ Pos						
24	193	1	▲192	63	▲130	1	▲192	40	▲153
106	68	1	▲67	103	▼-35	107	▼-39	163	▼-95
112	116	1	▲115	177	▼-61	140	▼-24	134	▼-18
163	1	1	■0	76	▼-75	103	▼-102	184	▼-183

Δ Position is the difference between the Original DEA ranking position and the ranking of the considered simulation.

While not all of the utilities experiment these dramatic changes, the robustness of the ranking (taking into account the quality of the data) is extremely weak and could easily be disputed by any of the regulated companies. The variability in the ranking has been summarized in Figure 2. This histogram shows utility groups according to the percentage of variation in their ranking (expressed as the positions changed with respect to the total number of companies). Only 2% of utilities maintain their position in all 81 DEA simulations, and only a 4% vary their position less than a 10%. However, around 60% of utilities see how their position in the ranking changes more than a 50%. Furthermore, there is an 8% of utilities varying their position more than a 90%.

Figure 2: Maximum variation in the ranking expressed in % for all utilities



This kind of variability would have a critical impact in a regulatory environment, as the consequences faced by a utility would change widely from being in the frontier (and therefore be considered a benchmark) or being in the last positions of the ranking. From a regulatory perspective, this kind of error in the assessment is serious enough to render any potential conclusions meaningless. After all, these relative positions are used in a yardstick competition environment to support key decisions affecting the economic part of the regulation.

Once again, it must be stressed that this variability is not a consequence of the utility modifying its performance or the context changing. They are simply reflecting the uncertainty in the data provided by the utilities, an uncertainty declared by the utilities as part of their data submission and subject to auditing from the regulator.

Considering the severe conclusions that may be drawn from the preceding results, some further consideration on the data used may be necessary. As presented when discussing the sample, the 194 utilities are already a subset of

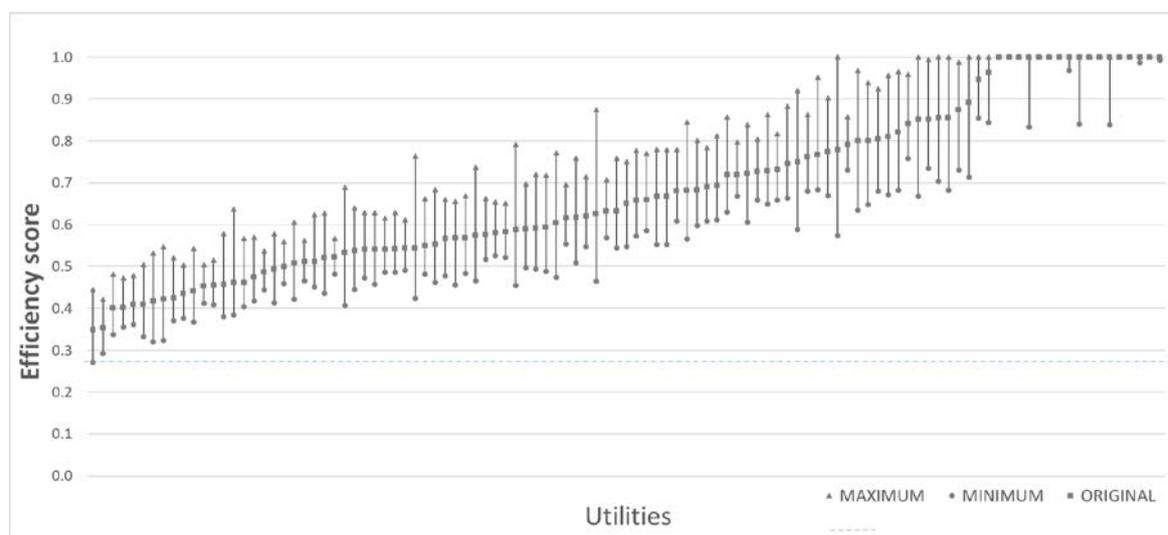
the 265 utilities reported by ERSAR; specifically, those that reported data quality information for all variables. However, a regulator must consider all regulated utilities in its region of mandate. Additionally, it must be reminded that since ERSAR is one of the few regulators requiring utilities to provide data quality information, we consider that regulated utilities in Portugal are sensitive to the need to produce good quality data (especially since all data are subject to audit by the regulator). This leads us to believe that the data quality of this particular set (European country, regulated, data quality reported and audited) can be considered to be good, or better than average.

However, in order to reinforce the conclusions, a second DEA simulation was run with a smaller cluster of utilities. The DMUs considered in this simulation were a subset of 108 utilities with uncertainties lower than 20% in all variables considered. In other words, the top 40% utilities (regarding the quality of their data). This simulation has been labelled as “Reduced sample”.

Figure 3 summarizes the results of the reduced sample simulation. As expected, the average amplitude has been reduced to a 17.40%. However, the maximum amplitude is still very large, 43%. These results, although better than the ones obtained in the complete sample simulation, are still far from producing reliable conclusions.

It can be deduced when comparing the complete simulation and the reduced sample simulation that DEA is a sensitive method when high uncertainties are considered. When the worst performers in terms of data quality have been removed from the sample, amplitude of range has decreased substantially.

Figure 3: Maximum, minimum and original DEA values for the reduced sample simulation



Respecting the number of efficient utilities, there are 18 efficient utilities in the Original simulation (without considering uncertainties). However, this value increases to 25 in the best scenario (7 utilities become efficient), and decreases to 12 in the worst scenario (6 utilities from the original scenario are declassified as efficient).

Table 3 shows an example of how the ranking vary in four different DEA simulations and four utilities. Comparing with the complete sample simulation, variation in ranking has decreased substantially. However, it is still significant. For instance, utility 53 varies from being in position 34 in the original DEA simulation to become efficient in another simulation.

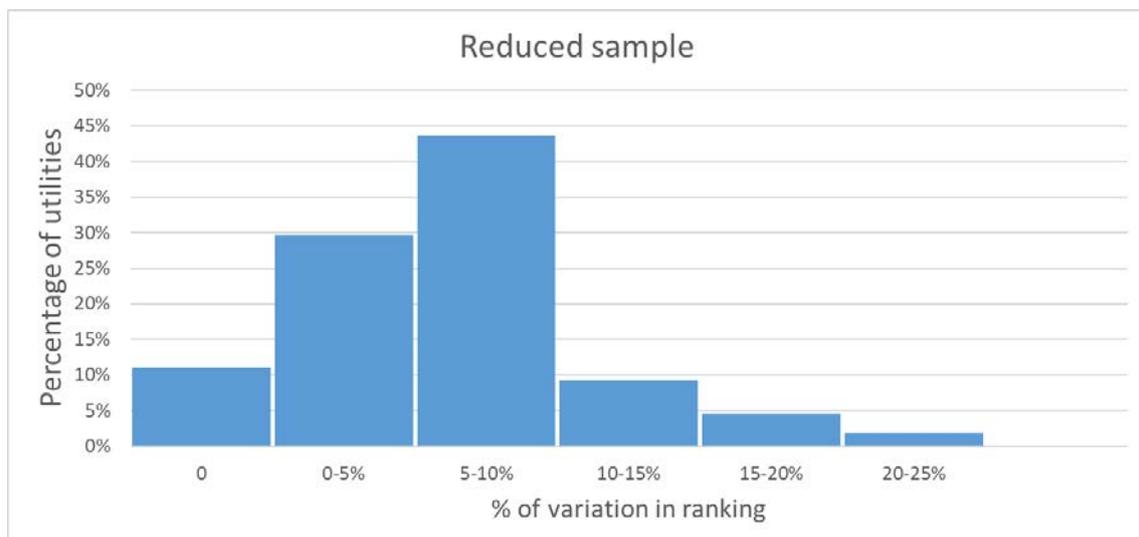
Table 3. Example of utilities variation in the ranking (reduced sample simulation)

ID Utility	ORIGINAL DEA Position	Simulation 6		Simulation 18		Simulation 24		Simulation 44	
	Ranking position	Ranking position	Δ Pos						
32	1	20	▼ -19	1	▲ 0	20	▼ -19	1	▲ 0
42	94	95	▼ -1	72	▲ 22	96	▼ -2	79	▲ 15
53	34	48	▼ -14	24	▲ 10	50	▼ -16	1	▲ 33
55	58	73	▼ -15	52	▲ 6	78	▼ -20	37	▲ 21

Δ Position is the difference between the Original DEA ranking position and the ranking of the considered simulation.

Figure 4 summarizes the variability in the ranking in the reduced sample simulation. In this case, an 84% of utilities vary their position between 0 and 10%. From them, only an 11% maintain their position and a 30% vary between 0 and 5%. However, there is still a 16% of utilities varying their position in the ranking between a 10 and a 25%.

Figure 4: Maximum variation in the ranking expressed in % for all utilities (reduced sample simulation)



As can be deduced from the presented results, uncertainties have an important impact on DEA results. The conclusions obtained when ignoring data quality, have shown little consistency once the potential variability of real data is considered.

The method shows greater consistency when better data are used. However, poor data quality is inherent to the water sector. Improving quality of data can require a significant amount of resources and a regulator can only promote better data quality, but not guarantee it.

Finally, this study shows that data quality needs to be considered an integral part of data collection for regulation. The validity of any method can only be tested if data quality information is collected and audited. Doing otherwise can only lead to doubtful conclusions when assessing the performance of competing utilities in a regulated environment.

5. Conclusions

Water services are usually provided in a natural monopoly market that requires to be regulated. It is the role of the regulator to guarantee that users are able to access services of a suitable quality, while their rights as consumers are protected with reasonable tariffs and a fair and open competition between service providers. Additionally information needs to be collected and presented in a transparent way.

Yardstick competition, as acknowledged since its inception, requires to model costs to eliminate context. DEA is the most popular of the non-parametric frontier methods used to assess efficiency, and as such has been extensively discussed and used for regulatory purposes. Until now, despite recommendations found in ISO standards (ISO 24510:2007, ISO 24511:2007, ISO 24512:2007) and the IWA manuals of best practices, data quality has not been a significant part of the discussion in regulation, especially when determining how suitable a method was to model costs or assess efficiency.

This paper presented a study to determine how sensitive DEA results are to the uncertainties that are inherently present in water sector data. Using public, audited data from a regulated environment, the results show that DEA is very sensitive to data quality and presents very inconclusive results when several scenarios are considered with the potential variation that data may present due to its uncertainty. In a regulatory environment, with a yardstick competition approach, a significant portion of the utilities considered could experiment significant differences in their perceived efficiency depending on pure chance.

Given the nature and origin of the studied utilities, it may be concluded that this uncertainty and, hence, its impact in DEA results, could be virtually extrapolated to any country in the world. DEA would work better with better quality of data, but the data quality for the whole sector in any other country can hardly be expected to be better than the uncertainty declared by the top utilities in Portugal.

Consequently, further research is needed to incorporate data quality in cost modelling and efficiency assessment. It remains to be seen if, with the adequate amendments, DEA will be an adequate methodology to assess comparative efficiency in the water sector with the quality of the currently available data.

In any case, based on the results of this study, it would be advisable for policy makers and regulators of water services to consider data quality as a substantial component of their efforts to assess efficiency. Such information should be used to carry out sensitivity analyses similar to the one presented here. This would prevent making critical regulatory decisions based on methodologies that could present significant uncertainties derived from poor data quality.

References

- Abbott, M., Cohen, B., 2009. Productivity and efficiency in the water industry. *Utilities Policy*, 17(3–4), pp.233–244. Available at: <http://dx.doi.org/10.1016/j.jup.2009.05.001>.
- Alegre, H., Baptista, J., Cabrera Jr, E., Cubillo, F., Duarte, P., Hirner, W., Merkel, W. and Parena, R., 2016. Performance Indicators for Water Supply Services: Third Edition. London: IWA Publishing. ISBN: 9781780406329
- Ananda, J., 2014. Evaluating the performance of urban water utilities: robust nonparametric approach. *Journal of Water Resources Planning & Management*, 140(9), 04014021.
- Arregui, F., Cabrera Jr., E., Cobacho, R., 2007. *Integrated Water Meter Management* 1st ed., London: IWA Publishing.
- Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30 (9), 1078-1092.
- Bhardwaj, B., Kaur, J., Kumar, A., 2018. A new approach for solving CCR data envelopment analysis model under uncertainty. *Studies in Fuzziness and Soft Computing*, 357, 319-343.
- Bonilla, M., Casasús, T., Medal, A., Sala, R., 2004. An efficiency analysis with tolerance of the Spanish port system. *International Journal of Transport Economics*, 31 (3), 379-400.
- Boscá, J.E., Liern, V., Sala, R., Martínez, A., 2011. Ranking decision making units by means of soft computing DEA models. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 19 (1), 115-134.
- Byrnes, J., Crase, L., Dollery, B. and Villano, R., 2010. The relative economic

- efficiency of urban water utilities in regional New South Wales and Victoria. *Resource and Energy Economics*. Volume 32, Issue 3, pp 439-455. <https://doi.org/10.1016/j.reseneeco.2009.08.001>
- Cabrera, E., Cabrera Jr., E., 2016. *Regulation of Urban Water Services. An Overview* 1st ed., London: IWA Publishing.
- Cooper, W.W., Seiford, L.M., Zhu, J., 2011. *Handbook on Data Envelopment Analysis*, Available at: <http://link.springer.com/10.1007/978-1-4419-6151-8>.
- Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2 (6), 429-444.
- De Witte, K., Marques, R.C., 2010a. Designing performance incentives, an international benchmark study in the water sector. *CEJOR* 18, 189e220. 189–220. <https://doi.org/10.1007/s10100-009-0108-0>
- De Witte, K., Marques, R.C., 2010b. Incorporating heterogeneity in non parametric models: a methodological comparison. *International Journal of Operational Research*, 9(2), 188–204.
- DiLorenzo T. J., 1996. The myth of natural monopoly. *The review of Austrian economics* Vol. 9, No. 2., pp 43-58. ISSN 0889-3047
- Dong, X., Zhang, X., Zeng, S., 2017. Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. *Water Research*, 112, 195-207.
- Dyson, R.G., Shale, E.A., 2010. Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society*, 61 (1), 25-34.
- ERSAR, 2017. Edições anuais do RASARP [Annual RASARP Editions]. www.ersar.pt. Available at: <http://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx> [Accessed November 10, 2017].
- ERSAR, 2015. ERSAR – Relatório Anual dos Serviços de Águas e Resíduos em Portugal. www.ersar.pt. Available at: <http://www.esgra.pt/ersar-relatorio-anual-dos-servicos-de-aguas-e-residuos-em-portugal/> [Accessed June 16, 2017].
- ERSAR, 2011. Guia de Avaliação da qualidade dos serviços se Águas E Resíduos prestados aos utilizadores. Publishing: Water and Waste Services Regulation Authority (ERSAR). National Laboratory for Civil Engineering (LNEC). ISBN: 978-989-8360-11-3
- Eslami, R., Khodabakhshi, M., Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Khoveyni, M., 2012. Estimating most productive scale size with imprecise-chance constrained input-output orientation model in data envelopment analysis. *Computers and Industrial Engineering*, 63(1), 254-261.
- Ferro, G., Lentini, E., Mercadier, A. and Romero, C., 2014. Efficiency in Brazil's water and sanitation sector and its relationship with regional provision, property and the independence of operators. *Utilities Policy*, 28, pp.42-51..
- Guerrini, A., Romano, G., Leardini, C., Martini, M., 2015. Measuring the efficiency of wastewater services through data envelopment analysis. *Water Science and Technology*, 71(12), 1845-1851.
- ISO 24510:2007: Service activities relating to drinking water supply systems and wastewater systems-- Guidelines for the assessment and for the improvement of the service to users. International Organization for

- Standardization, Geneva, Switzerland. Available at: <https://www.iso.org/standard/37246.html>
- ISO 24511:2007: Activities relating to drinking water and wastewater services -- Guidelines for the management of wastewater utilities and for the assessment of wastewater services. International Organization for Standardization, Geneva, Switzerland. Available at: <https://www.iso.org/standard/37247.html>
- ISO 24512:2007: Activities relating to drinking water and wastewater services -- Guidelines for the management of drinking water utilities and for the assessment of drinking water services. International Organization for Standardization, Geneva, Switzerland. Available at: <https://www.iso.org/standard/37248.html>
- ISO 4064-1:2014: Water meters for cold potable water and hot water -- Part 1: Metrological and technical requirements. International Organization for Standardization, Geneva, Switzerland. Available at: <https://www.iso.org/standard/55371.html>
- IWA - International Water Association, 2015. The Lisbon Charter.
- Joskow, P. L., 2007. Regulation of natural monopolies. *Handbook of Law and Economics*, Volume 2. ISBN 978-0-444-53120-9.
- Kao, C., Liu, S. T., 2014. Measuring performance improvement of Taiwanese commercial banks under uncertainty. *European Journal of Operational Research*, 235 (3), 755-764.
- Marques, R.C., Berg, S., Yane, S., 2014. Nonparametric benchmarking of Japanese water utilities: Institutional and environmental factors affecting efficiency. *Journal of Water Resources Planning and Management*, 140 (5), 562-571.
- Molinos-Senante, M., Donoso, G., Sala-Garrido, R., 2016. Assessing the efficiency of Chilean water and sewerage companies accounting for uncertainty. *Environmental Science and Policy*, 61, 116-123.
- Molinos-Senante, M., Donoso, G., Sala-Garrido, R., Villegas, A., 2018. Benchmarking the efficiency of the Chilean water and sewerage companies: a double-bootstrap approach. *Environmental Science and Pollution Research*, 25, 8432-8440.
- Picazo-Tadeo, A.J., Sáez-Fernández, F.J., González-Gómez, F., 2008. Does service quality matter in measuring the performance of water utilities? *Utilities Policy*, 16(1), pp.30–38.
- Pinto, F.S., Simões, P., Marques, R.C., 2016. Water services performance: do operational environment and quality factors count? *Urban Water Journal*, 9006(December), pp.1–9. Available at: <https://www.tandfonline.com/doi/full/10.1080/1573062X.2016.1254254>.
- Posner, R., 1969. Natural Monopoly and its Regulation. *Stanford Law Review*, 21(3), pp.548–643. Available at: <http://www.jstor.org/stable/1227624>.
- Romano, G., Guerrini, A., 2011. Measuring and comparing the efficiency of water utility companies: A data envelopment analysis approach. *Utilities Policy*, Vol 19, Issue 3, pp 202-209. <https://doi.org/10.1016/j.jup.2011.05.005>
- Sawkins J.W., 1995. Yardstick competition in the English and Welsh water industry. Fiction or reality? *Utilities Policy*. Vol. 5, No.1, pp 27- 36. 0957-1787/95
- See, K.F., 2015. Exploring and analysing sources of technical efficiency in

- water supply services: Some evidence from Southeast Asian public water utilities. *Water Resources and Economics*, 9, 23-44.
- Shleifer, A., 1985. A theory of yardstick competition. *The Rand Journal of Economics*. Vol 16, No. 3, pp 319-327
- Tupper, H.C., Resende, M., 2004. Efficiency and regulatory issues in the Brazilian water and sewage sector: An empirical study. *Utilities Policy*, 12(1), pp.29–40.
- United Nations, 2010. The human right to water and sanitation. Resolution adopted by the General Assembly on 28 July 2010. A/RES/64/292. <http://daccess-ods.un.org/access.nsf/Get?Open&DS=A/RES/64/292&Lang=E>
- Walter, M., Cullmann, A., von Hirschhausen, C., Wand, R. and Zschille, M., 2009. Quo vadis efficiency analysis of water distribution? A comparative literature review. *Utilities Policy*, 17(3–4), pp.225–232. Available at: <http://dx.doi.org/10.1016/j.jup.2009.05.002>
- Wang, K., Wei, Y.-M., Zhang, X., 2012. A comparative analysis of China's regional energy and emission performance: Which is the better way to deal with undesirable outputs? *Energy Policy*, 46, 574-584.
- Worthington, A.C., 2014. A review of frontier approaches to efficiency and productivity measurement in urban water utilities. *Urban Water Journal*, 11(1), 55-73.